VARIAZIONI: Collaborative Authoring of Localized Cultural Heritage Contents over the Next Generation of Mashup Web Services

Carlos A. Iglesias
Germinus XXI (Grupo Gesfor)
Avda. Manoteras, 32
28050 Madrid - Spain
cif@germinus.com

Marta Sánchez
Fundación Albéniz
C/ Barquillo, 6
28004 Madrid – Spain
msanchez@albeniz.com

Francesco Spadoni
Rigel Engineering s.r.l
Via Spagna, 10
57017 Guasticce - Italy
spadoni@rigel.li.it

Abstract

This paper presents developments and results from the VARIAZIONI[1] project, funded by the Commission of the European Union under the eContentPlus program. The project aims to develop and validate an open collaborative infrastructure and methodology for authoring, distribution and delivery of European Musical Contents. VARIAZIONI approach encompasses the development of a collaborative platform where non technical users from cultural institutions can easily catalogue their existing digital assets, comprising multimedia master classes, audio/video concerts, images, scanned documents, etc., and create new contents from the existing contents of other institutions, taking into account IPR issues. We aim to demonstrate the feasibility of reusing and integrating contents in different applications, like websites, e-courses, collaborative portals, and new distribution channels, through the usage of friendly interfaces and intelligent techniques for reducing cataloguing effort. As a result, existing contents will be more accessible and enriched through its reuse and inclusion of internet resources.

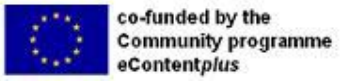
1. Introduction

The status of culture in the European society is constantly growing. Music is one of the cultural domains where the importance of the European heritage is essential. During the last years, many European projects were working in the development of contents and tools to attend the increasing demand in musical products and services. Among them, and since April 2004, a partnership of the major musical institutions in Europe have been working in the development of the e-Content HARMOS project, whose goal has been to preserve and reuse information produced by musical institutions in the area of Cultural Heritage., a partnership formed for Royal College of Music, Lithuanian Academy of Music, Koninklijk Conservatorium Brussels, Staatliche Hochschule für Musik und Darstellende Kunst Stuttgart, Escola Superior de Música de Catalunya, Escola Superior de Música do Porto, Orbiteam, Universitat Pompeu Fabra and Germinus XXI, worked together, under the coordination of Fundación Albéniz in the production of a digital collection and a technological platform to disseminate and preserve these valuable contents.

At the same time, many European projects as MUSICNETWORK, WEDELMUSIC and AXMEDIS were working in research activities by adopting architecture of participation and involving users from the very stage of content creation, classification and transformation.

Using the results and expertise of all these European projects, VARIAZIONI has the main goal of promoting the reuse and enrichment of existing musical digital assets in new applications and new distribution channels using Web 2.0 principles: consider the web as

[1] This Project has received funding from the eContentPlus programme of the European Community. The authors are solely responsible of this communication nad it does not represent the opinion of the European Community. The European Community is not responsible of any use that might be made of information contained herein.



0-7695-3030-3/07 $25.00 © 2007 IEEE
DOI 10.1109/AXMEDIS.2007.27

a platform (collection of "web services" to integrate together to build different systems) and to build an architecture of participation to the long tail and not just the head. To reach this aim, VARIAZIONI project proposes:

- The development of a collaborative platform where non technical users from cultural institutions can easily catalogue their existing digital assets.
- Re-use contents from the existing contents of other institutions taking into account IPR / DRM issues.
- To test and demonstrate the feasibility of reusing and integrating contents in different applications and new distribution channels.
- To make existing contents more accessible and enriched through its reuse and inclusion of Internet resources.

The project will develop a pilot where major European Musical Institutions will develop and integrate existing musical contents and will demonstrate its reuse (through the development of courses, teaching material, articles, thematic webs, …). In addition, the pilot will show how these contents can be enriched through the usage of automatic semantic tagging with audio processing and web 2.0 technologies. The project leverages the results and communities arose from eContent HARMOS project and IST AXMEDIS project.

2. Objectives and approaches

The name of the project, VARIAZIONI, comes from the Italian name of one of the earliest musical forms: variations. A theme or melody is presented in a straightforward manner and then repeated as often as the composer wishes but each time it is varied in one or more ways. From a technical perspective, we could describe this process as the process where a musical content (the theme) is reused and enriched in every variation.

The process of content enrichment in VARIAZIONI is based on **content aspects** which are filled in automatically and collaboratively. An aspect is a set of related metadata about a perspective of a content (dynamics, body position, technical aspects ...). In the process of reusing content for different contexts, content aspects (metadata) are added. Thus, content is enriched with metadata relevant for this new context. Let us provide a user case for developing this important concept for enriching contents through their "variazioni".

Use Case

A Musical Institution records a concert of a duo of mandolin and piano. The musical institution adds some metadata about the video (event, room, interpreters, date, instruments, etc.) and makes available the video in the VARIAZIONI demonstrator.

Let's say a specialist in organology wants to reuse the video for an article about the origin of the mandolin. The specialist accesses VARIAZIONI demonstrator portal, selects a new aspect of the content (organology) for tagging this new information and enriches the content (etymology of 'mandolin', precedent instruments, etc.). He writes then the article "Origins of the Neopolitan mandolin". The original content has been enriched with this new aspect. These metadata enrich also the rest of contents catalogued with the term 'mandolin', since the metadata is semantically enriched (ontology). The article is published in the community web site and other communities have the article available for its reuse and / or enrichment. In addition, users add their own defined tags for this article. The specialist can review the user tagging and add some of the most popular tags to its original article, increasing the quality of its tagging.

Now a luthier wants to publish on a web site a lesson on lutherie, and selects the same video. He produces a new "variazioni" of the content with this new aspect, lutherie. Here the relevant metadata included in this new aspect are the luthier of the mandolin, the woods used in its construction, etc. Then, he tags the metadata of the new aspect, generating an enriched content and reuses this enriched content in a new content, a lesson about the description of the art of making mandolins by this luthier. In order to reduce tagging effort, an automatic tagging assistant has helped to fill the biography of the luthier, and some relevant biographical metadata that has been automatically extracted from Wikipedia and tagged.

Later a cultural institution selects the video for an announcement of a concert by the same duo with the same programme. They add a new aspect, cultural event, and tag the metadata about the programme of the video, since it the initial tagging was not so complete.

Then include in the announcement (a new content, concert announcement), the concert place, date, etc. and link the video. Automatically, a summarisation of the pieces of the video is created and links to the similar concerts in the same area.

The community approach facilitates tagging distributed resources and enriching the tagging through their "variazioni" and appropriate tools for manual and automatic tagging.

The project VARIAZIONI has the goal of improving Musical Metadata Tagging through:

- the definition of a Common European Musical Metadata and
- a Collaborative Tagging and Enrichment Process based on Tools that combine automatic and manual tagging.

One of the most innovative aspects of VARIAZIONI is that is based on **communities**. Internet has shown the power of communities. VARIAZIONI proposes to develop a platform in which protagonist is transferred to the users, who maintain and create the contents. In VARIAZIONI, the process on enrichment is based on adding new metadata in the process of reusing (adapting) contents for new contexts or even customising content metadata by a user.

The project VARIAZIONI will setup a collaborative platform, VARIAZIONI Content Enrichment Portal, where users can contribute to the creation, enrichment and cataloguing of musical contents.

VARIAZIONI Content Enrichment Portal aims to develop and validate an open infrastructure and methodology for distribution and delivery of European regional cultural content, mainly based on music content. The project will provide a paradigm shift in how the content is created, transformed, adapted and consumed, by adopting an architecture of participation and involving users from the very early stage of content creation, classification and transformation [6].

The project is built on top of relevant results of previously research activities in European R&D Programs (MUSICNETWORK [6], WEDELMUSIC [7], SIMAC [8]), eContent projects such as HARMOS [1] and in particular on the ongoing project AXMEDIS [11] based on a technology platform, with DRM support and tools for production, protection and distribution of content.

Using the results and expertise of all these European projects, VARIAZIONI has the main goal of improving Musical Metadata Tagging and its quality, through the promotion of the reuse and enrichment of existing musical digital assets in new applications and new distribution channels.

The main issue of the HARMOS project was the preservation of educational contents from the great European maestros thought the recording in audiovisual format of their classes. The necessity to work with the audiovisual archives of the participating institutions made us aware both of the value of the materials that all of these institution have in their archives and of the difficulty to undertake common projects due the high variety of formats, cataloguing tools and even preservations conditions. The work developed by MUSICNETWORK or WEDELMUSIC projects showed the importance of this cultural heritage. The huge dimension of the work ahead made relevant the need of a new project in order to preserve and make accessible this cultural content in a very usable way through the definition and application of a Common European Musical Metadata.

The main objectives of the project are the following:

- O1. Define a common metadata for European Musical Digital assets
- O2. Automatic enrichment of contents with available Mashup web services and automatic tagging systems
- O3. Provide a community driven open infrastructure for content enrichment, re-use and secure distribution
- O4. Provide a platform and rights model for sustainability of content enrichment
- O5. Leverage user participation to enrich content and content metadata
- O6. Promote the development of user communities for continuous enrichment of contents
- O7. Continuous improvement of contents and services through user behaviour real-time monitoring

3. Architecture

From a technological point of view, VARIAZIONI architecture is based on four main components as depicted in Figure 1, which are described below:

- a standard platform (JSR-168) based on portlets (Xpression Portal)
- a standard Content Management System (Java Content Repository, JSR-170) (Xpression CMS)
- set of tools for protection and production of protected contents based on MPEG-21 (Axmedis);
- set of tools for semantic content enrichment (multifaceted User Interface and wizards for folksonomies, mashups-based mechanisms for metadata extraction from external sources) [10].

The proposed architecture of VARIAZIONI exhibits the following qualities:

- Standardisation. The architecture is based on the well known specifications (Portlets [2], Java Content Repository [3] and MPEG-21[4][5])
- Integration of legacy system. The portlet infrastructure is used for integrating existing

cataloguing tools. This facility has been used for integrating results from HARMOS as a demonstration of this capability,

- Integration of external annotation systems. The portal infrastructure provides a REST layer for managing content metadata. This layer is used by the semantic enrichment tool, which adds metadata as a result of audio analysis of the contents, smart clipping of resources available in Internet, or social derived taxonomies (folksonomies) according to some of the web2.0 paradigms [9],
- Usage of mainstream content management system technology for managing the life cycle of the contents, integrated with a collaborative system, that allows its re-purposing in the web sites associated to every community [10],
- Integration with a robust standards-based AXMEDIS platform for content protection and delivery [12]. AXMEDIS follows a grid-based architecture for processing high volume of contents and is based on the standard MPEG-21. It provides (i) content transformation and packing for producing content in a suitable format for delivering to different devices, and (ii) content protection and DRM support.

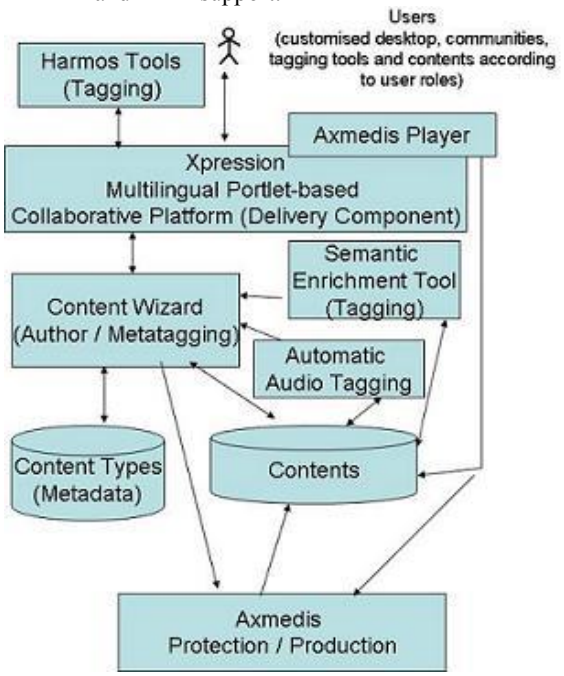


Figure 2. VARIAZIONI Architecture

4. Collaborative Enrichment

VARIAZIONI proposal for annotation is based on social networking. A portlet-based collaborative platform has been set up where users can share contents in the communities they belong to.

At this stage of the project, there is a first definition of Musical Metadata for Variazioni project, and content wizards has been created for its annotation, as shown in Figure 3.

Figure 3. Collaborative authoring

Contents from HARMOS project have been exported as SCORM zips, and imported into the content management system.

A first prototype of content enrichment through social tagging has been developed as shown in Figure 4.

Figure 4. Social tagging of resources

5. Folksonomies and web2.0 concepts

VARIAZIONI adopts and refines emerging collaborative practices for content enrichment based on web2.0 concepts, leveraging the use of folksonomies , focusing on user participation, and exploiting the architecture of the web as a platform.

In particular, we developed tools and facilities for exploiting socially derived taxonomies (i.e. folksonomies). This classification schema has proven to

be very accurate when communities tag the same resources, and in addition it provides feedback for improving quality of tagged resources.

Such tools enable the participation of user communities in the classification of existing content they are interested in, as well as content they create/integrate. Hence, the tools developed support the creation of user generated tags (in form of folksonomies) which can complement and enrich the existing metadata. The tools consist of a rich multifaceted user interface (UI) adapted to the specific VARIAZIONI content and metadata, wizards for tags selection and insertion, powerful functions for simultaneous tagging operations on multiple content objects, support for quality assurance mechanisms (support for evaluation, revisions and modifications of tags).

VARIAZIONI tools allow users to tag content objects with a descriptive word, expressing a characteristic of the content or associated meaning. They represent folksonomies as a tag cloud, which displays the most popular tags;

When a user creates a tag for a specific content object, the tag is stored in the database and associated with that object (by its ID). The system keeps track of all the tags that users have entered and the number of times that they have entered the same tag.

Each tag is visualized with a font size based on the popularity of that tag. This allows users to browse the content by way of a user-driven categorization of that content. When a user clicks on one of the tags in the tag cloud, the application retrieves a list of the associated contents. searching by tags (folksonomies) or by metadata.

6. User focus and participation

The development of VARIAZIONI components mainly targeted to end-users, like the folksonomies tools or the wizards, followed an iterative approach based on Unified Process with the aim to gather early feedbacks from users, thus ensuring an improved chance of meeting their requirements and needs.

The tools were developed with the aim to facilitate and stimulate users contribution (i.e. high usability standards, powerful functions for batch metatagging operations on multiple objects, simple yet appealing look and feel).

Like most advanced digital users, VARIAZIONI users are increasingly interested in accessing all the aspects of a digital content, like user-generated video, photos, podcasts, music, games and more. They want access to all available data, all in real-time. VARIAZIONI is leveraging the users themselves to help organize the content and make it accessible and searcheable, using the “wisdom of the crowds”. This is being enabled by the use of tags and tagging, which helps to unify and organize the rapidly expanding world of user-generated media, as described in Section 6. As VARIAZIONI CEP support the use of tags, it is able to assemble collections of social media based on the interests of VARIAZIONI users.

Significant effort went in making the user interface simpler, clean and more intuitive, doing user testing, performing validation sessions and listening to VARIAZIONI users, collecting and prioritizing what they wanted, liked and disliked.

7. Conclusions and future work

The project VARIAZIONI is addressing content enrichment taking advantage of industrial practice, research progress and emerging web 2.0 technologies, which will provide a basis for its commercialisation and sustainability.

The involvement of musical institutions in their conception is essential for its acceptance in the musical community.

Next steps in the project are the usage of existing tools for tagging musical resources, the integration of AXMEDIS platform, and the integration of semantic enrichment tools.

In addition, the content enrichment portal will be used for the validation of the proposed approach to content enrichment by directly involving the user communities in the process. To this aim, mechanisms are being developed and integrated into the portal, allowing to monitor and analyse user behaviour and to measure important quantitative indicators (navigation and search paths, usage of features, some usability metrics) to understand the relevance and accuracy of classification and tagging. In addition, the content enrichment portal will provide support for preparing and submitting questionnaires to users to gather qualitative metrics (related to visibility, accessibility and searchability).

8. Acknowledgments

The authors would like to thank the contributions from the partners of VARIAZIONI, and the constructive and helpful comments from Spyridon Pilos along the project. In particular, the authors would like to thank specially to Almudena Albertos, Paolo Nesi, Emilia Gómez, Piero Alcamo and Daniel Molina.

9. References

[1] Carlos A Iglesias, Marta Sánchez, A. Guibert, M. J. Guibert and E. Gómez, A Multilingual Web-based Educational System for Professional Musicians, Proceedings of the IV International Conference on Multimedia and Information and Communication Technologies in Education, M-ICTE, Seville, 2006.

[2] Java Portlet Specification (JSR-168), en http://jcp.org/en/jsr/detail?id=168 .

[3] Java Content Repository (JSR-170), JSR-170. Content Repository for Java Technology API, Day Software, Junio 2005.

[4] MPEG-21 Information Technology – Multimedia Framework – MPEG21 – International Standard ISO/IEC 21000-2, 2nd Edition, 2005.

[5] Michel Girer, Bernhard Günther, Martin Schmucker, Francesco Spadoni, "MPEG-21 and Music Notation Applications", Proceedings of the 4th IEEE Conference on Web Delivering of Music Scores, Barcelona (E), September 13-14, 2004.

[6] Francesco Spadoni., Riccardo Lenzi Martin Schmucker, Distribution of Coded Music. In Mitolo M., Nesi P., Ng K. (Ed.). Proceedings of the 5th MUSICNETWORK Open Workshop, Vienna (A), July 2005.

[7] Riccardo Lenzi, Francesco Spadoni, Alfonso Tatarelli, "An Application for Multimedia Content Management and On-line Distribution", Proceedings of the 2nd IEEE Conference on Web Delivering of Music Scores, Darmstadt (D), December 9-11, 2002.

[8] Semantic Interaction with Music Audio Contents, SIMAC Project, http://www.semanticaudio.org/ .

[9] Tim O'Reilly, "What is web 2.0 – the wisdom of the crowds", http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html

[10] Dan Bricklin, "The Cornucopia of the Commons: how to get volunteer labor", http://www.bricklin.com/cornucopia.htm

[11] Ng, K., Ong, B., Neagle, R., Ebinger, P., Schmucker, M., Bruno, I., Nesi, P., "Axmedis framework for programme and publication and on-demand production", First International Conference on Automated Production of Cross Media Content for Multi-channel Distribution, December 2005.

[12] Paolo Nesi, Davide Rogai, Andrea Vallotti, "A protection processor for MPEG-21 players", IEEE International Conference on Multimedia & Expo (ICME 2006), Toronto, Canada, July, 2006.